XIXth National Conference
CIRCUIT THEORY AND ELECTRONIC NETWORKS
Kraków–Krynica, Poland, October 23th–26th, 1996

Włodzimierz Greblicki
Zygmunt Hasiewicz
Institute of Engineering Cybernetics
Technical University of Wrocław
Wybrzeże S. Wyspiańskiego 27, 51-370 Wrocław

Identification of Non-linear Systems Corrupted by Colored Noise

ABSTRACT

Memoryless, Hammerstein, and Wiener systems are identified. The paper focuses on recovering their non-linear parts. The systems are driven by random signals and are disturbed by colored additive random noise. The *a priori* information about the systems is non-parametric in nature. To estimate non-linear characteristics of examined systems, an adequate non-parametric algorithm is proposed. Its convergence is shown and convergence rate is given.

1. INTRODUCTION

The parametric identification of linear systems disturbed by white random noise is well elaborated in the literature. Problems arise when the noise is colored. In such situations, identification algorithms do not behave well and require severe modifications. In this paper, we examine the non-parametric approach, which means that our *a priori* information about the system to be identified is much smaller than in parametric inference. It makes the problems closer to those encountered in applications. We propose a class of algorithms to estimate non-linearities in memoryless, Hammerstein, and Wiener systems. We show that our algorithms successfully recover non-linearities for white and colored noise as well.

2. MEMORYLESS SYSTEM IDENTIFICATION

We begin our considerations with a very simple system. The system, shown in Fig. 1, is namely memoryless and has a characteristic m. Its input is a stationary white random noise $\{U_n; n = ..., -1, 0, 1, 2, ...\}$ with finite variance. The probability density of U_n exists and is denoted by f. The system is disturbed by stationary random noise $\{Z_n; n = ..., -1, 0, 1, 2, ...\}$. The noise itself is an output of an asymptotically stable linear system driven by stationary random noise $\{\xi_n; n = ..., -1, 0, 1, 2, ...\}$ with zero mean and finite variance. Processes $\{U_n\}$ and $\{\xi_n\}$ are mutually independent. The non-linear characteristic of the system is completely unknown and our problem is to estimate the true characteristic from input-output observations (U_0,Y_0), (U_1,Y_1), ..., (U_n, Y_n).

Observe that *a priori* information concerning m is extremely poor since the class of admissible (at present: all possible) characteristics is so ample that no parametric representation of the class is known. Thus our identification problem is, in fact, non-parametric. Moreover, and this is in contrast to the problems which have already been examined in the literature - see e.g. [7], the noise $\{Z_n\}$ is correlated due to the fact that it is the output of a dynamic system.

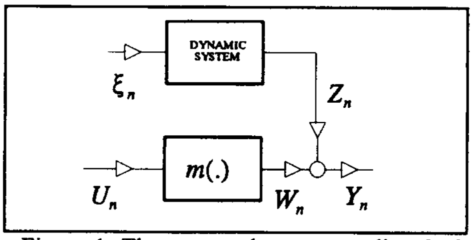

Figure 1. The memoryless system disturbed by noise Z_n

Let us notice that

$$E\{Y_n | U_n = u\} = m(u), \tag{2.1}$$

i.e., that m is a regression function. To estimate the unknown characteristic, we apply the following regression estimate:

$$\hat{m}(u) = \sum_{i=1}^{n} Y_i K(\frac{u - U_i}{h(n)}) \Big/ \sum_{i=1}^{n} K(\frac{u - U_i}{h(n)}), \tag{2.2}$$

where K is a non-negative kernel function, and $\{h(n)\}$ a positive number sequence. The kernel satisfies the following restrictions:

$$\sup_{u \in (-\infty, \infty)} K(u) < \infty, \quad \int_{-\infty}^{\infty} K(u) du < \infty, \quad |u| K(u) \to 0 \text{ as } |u| \to \infty. \tag{2.3}$$

In turn, the positive number sequence is such that

$$h(n) \to 0 \text{ as } n \to \infty, \tag{2.4}$$

$$nh(n) \to \infty \text{ as } n \to \infty. \tag{2.5}$$

Theorem 1. Let $E\{m^2(U_0)\} < \infty$. Let the non-negative Borel kernel satisfy (2.3) and let the positive number sequence $\{h(n)\}$ fulfil (2.4)-(2.5). Then

$$\hat{m}(u) \to m(u) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point u at which both f and m are continuous, and $f(u) > 0$.

Proof. Let $\hat{g}(u) = (1/nh(n))\Sigma_{i=1}^{n} Y_i K[(u - U_i)/h(n)]$, $\hat{f}(u) = (1/nh(n))\Sigma_{i=1}^{n} K[(u - U_i)/h(n)]$. Observe $\hat{m}(u) = \hat{g}(u)/\hat{f}(u)$. We have $E\hat{g}(u) = (1/h(n))E\{m(U_0)K[(u - U_0)/h(n)]\}$. By virtue of Lemma 2 in Appendix, and (2.4), we obtain

$$E\hat{g}(u) \to f(u)m(u) \int_{-\infty}^{\infty} K(v) dv \text{ as } n \to \infty. \tag{2.6}$$

In turn, $\text{var}[\hat{g}(u)] = V_1(u) + V_2(u)$, where

$$V_1(u) = \frac{1}{n^2 h^2(n)} \text{var}[\sum_{i=1}^{n} m(U_i) K(\frac{u - U_i}{h(n)})], \text{ and } V_2(u) = \frac{1}{n^2 h^2(n)} \text{var}[\sum_{i=1}^{n} Z_i K(\frac{u - U_i}{h(n)})],$$

respectively. Since $\{U_n\}$ is a white process,

$$V_1(u) = \frac{1}{nh^2(n)}\mathrm{var}[m(U_0)K(\frac{u-U_0}{h(n)})] \le (\kappa/nh(n))\frac{1}{h(n)}E\{m^2(U_0)K(\frac{u-U_0}{h(n)})\},$$

where $\kappa = \sup_u K(u)$. Applying Lemma 2 in Appendix, we find $V_1(u) = O(1/nh(n))$ as $n \to \infty$, at every point at which both f and m are continuous. Examining $V_2(u)$ is tougher since $\{Z_n\}$ is not white. We have

$$V_2(u) = \frac{1}{n^2h^2(n)}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{cov}[Z_iK(\frac{u-U_i}{h(n)}), Z_jK(\frac{u-U_j}{h(n)})].$$

Since $\{U_n\}$ and $\{Z_n\}$ are mutually independent and $EZ_n=0$, the covariance in the above expression equals or is smaller than

$$\mathrm{cov}[Z_i, Z_j]\,E\{K(\frac{u-U_i}{h(n)})K(\frac{u-U_j}{h(n)})\} \le \kappa\,\mathrm{cov}[Z_i, Z_j]\,E\{K(\frac{u-U_0}{h(n)})\},$$

with κ defined above. Therefore,

$$V_2(u) \le \frac{\kappa}{n^2h(n)}\frac{1}{h(n)}E\{K(\frac{u-U_0}{h(n)})\}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{cov}[Z_i, Z_j].$$

Invoking Lemma 1 in Appendix, we obtain the above quantity bounded by $c_1[1/nh(n)] \times (1/h(n))E\{K[(u-U_0)/h(n)]\}$, some c_1 independent of n. Applying Lemma 2 in Appendix, we find the quantity of order $O(1/nh(n))$ at every point at which f is continuous. Thus, $\mathrm{var}[\hat{g}(u)] = O(1/nh(n))$ at every continuity point of f. In this way, we have shown that

$$\hat{g}(u) \to f(u)m(u)\int_{-\infty}^{\infty}K(v)dv \quad \text{as} \quad n \to \infty \quad \text{in probability},$$

at every point u at which both f and m are continuous.

Since, using similar arguments, we can show that $\hat{f}(u)$ converges to $f(u)\int_{-\infty}^{\infty}K(v)dv$ (in probability) as n tends to infinity, the proof has been completed. □

It is possible to verify that for a suitably selected kernel and number sequence,

$$|\hat{m}(u) - m(u)| = O(n^{-2/5}) \text{ in probability as } n \to \infty,$$

at every point at which both m and f are twice differentiable.

We want to emphasize the fact that correlation of the noise does not decrease the convergence rate of (2.2), see e.g. [7] for suitable results in the white noise case. Observe, moreover, that compared to $n^{-1/2}$, i.e., the rate typical for parametric inference, our rate is quite good. It should be here remarked that in almost all papers devoted to recovering system characteristics from noisy data the *a priori* information is much greater than that assumed in this paper. Many authors assume for instance that m is a polynomial with known order. Then the (parametric) identification problem is just to estimate the polynomial coefficients (see, e.g., [1] and the papers cited therein). It seems that, in this respect, assuming poor *a priori* information about the system and considering the identification problem as a non-parametric one, we are closer to real situations.

We have assumed that the noise disturbing the system is not white but correlated. It leads to some difficulties in the proof of Theorem 1. On the other hand, however, the fact that the noise is colored does not cause any numerical complications. We want to stress strongly this nice property of our estimate. As far as parametric algorithms are concerned, correlation in the noise results both in greater theoretical and numerical problems. Identification algorithms should be also suitably modified.

3. HAMMERSTEIN SYSTEM IDENTIFICATION

In this section we consider a Hammerstein system shown in Fig. 2. The system consists of two subsystems connected in a cascade. The non-linear memoryless component has a characteristic m (i.e. $W_n = m(U_n)$) and the linear dynamic part is by assumption asymptotically stable. The system is driven by a stationary white random noise $\{U_n; n = ...-1, 0, 1, 2, ...\}$ and disturbed by zero mean stationary random noise $\{Z_n; n = ...-1, 0, 1, 2, ...\}$. Properties of the noise are the same as in Section 2. So is the *a priori* knowledge about the nonlinear characteristic. The goal is to identify both system components from input-output observations of the overall system (U_0,Y_0), (U_1,Y_1), ..., (U_n, Y_n). Here, we shall however focus our attention on the identification of the first subsystem only.

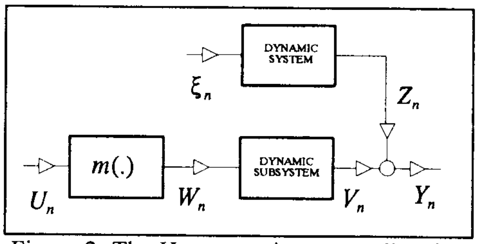

Figure 2. The Hammerstein system disturbed by noise Z_n

We begin our considerations with the following observation:

$$E\{Y_n | U_{n-1} = u\} = c_1 m(u), \tag{3.1}$$

where c_1 is a constant (we have assumed for simplicity that EW_n=0; this is the case for ,e.g., m odd and f even). Thus recovering $c_1 m(u)$ is equivalent to estimating the regression in (3.1). The fact that we can estimate the unknown characteristic only up to some multiplicative constant is a simple consequence of the cascade structure of the system. To recover m, we propose the following estimate:

$$\bar{m}(u) = \sum_{i=1}^{n} Y_{i+1} K\left(\frac{u - U_i}{h(n)}\right) \Bigg/ \sum_{i=1}^{n} K\left(\frac{u - U_i}{h(n)}\right), \tag{3.2}$$

where K and $h(n)$ are the same as in (2.2). Observe that the estimate is an appropriate modification of (2.2). Our next theorem will be given without a proof.

Theorem 2. Let m be odd and let $E\{m^2(U_0)\} < \infty$. Let f be even. Let the non-negative Borel kernel satisfy (2.3) and the positive number sequence $\{h(n)\}$ fulfil (2.4)-(2.5). Then

$$\bar{m}(u) \to c_1 m(u) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point u at which both f and m are continuous, and $f(u) > 0$.

For a suitably selected kernel and number sequence,

$$|\hat{m}(u) - c_1 m(u)| = O(n^{-2/5}) \text{ in probability as } n \to \infty,$$

at every point at which both m and f are twice differentiable. Observe that the convergence rate is the same as in Section 2.

4. WIENER SYSTEM IDENTIFICATION

In the Wiener system, Fig. 3, the linear dynamics is followed by the static non-linearity. The restrictions imposed on both subsystems, as well as the noise, are the same as in Section 3. The only difference is that U_n's have now, by assumption, a normal distribution with zero mean. The characteristic m can be of any form. We, however, assume that the characteristic is invertible, i.e., that m^{-1} exists. It can be then verified that

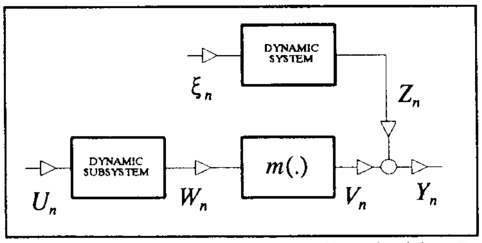

Figure 3. The Wiener system disturbed by Z_n

$$E\{U_n | Y_{n+1} = y\} = c_2 m^{-1}(y), \tag{4.1}$$

where c_2 is a constant. Thus, estimating the regression in (4.1), we can recover the inverse of m up to some multiplicative constant. We use the following algorithm:

$$\hat{m}(y) = \sum_{i=1}^{n} U_i K\left(\frac{y - Y_{i+1}}{h(n)}\right) \Bigg/ \sum_{i=1}^{n} K\left(\frac{y - Y_{i+1}}{h(n)}\right). \tag{4.2}$$

The estimate is an obvious modification of the estimate (2.2). The proof of the theorem given below will be presented in later papers.

Theorem 3. Let m be invertible and twice differentiable. Let the non-negative Borel kernel satisfy (2.3) and let, moreover, K be a Lipschitz function. Let the positive number sequence $\{h(n)\}$ fulfil (2.4) and let

$$nh^2(n) \to \infty \text{ as } n \to \infty. \tag{4.3}$$

Then

$$\hat{m}(y) \to c_2 m^{-1}(y) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point y at which the density of Y_n is positive.

By an appropriate selection of a kernel and number sequence, we can obtain

$$|\hat{m}(y) - c_2 m^{-1}(y)| = O(n^{-1/3}) \text{ in probability as } n \to \infty,$$

at every point at which both m and f have three derivatives. Observe that the convergence rate is now worse than that in Sections 2 and 3. Nevertheless, as it can be verified, it is the same for white and colored noise.

5. FINAL REMARKS

There exist kernel functions satisfying (2.3). One can choose, e.g., a rectangular kernel equal 1 or 0 according to $|u| \le 1$ or $|u| > 1$, respectively. Other kernels are, e.g., $\exp(-|u|)$, $1/(1+u^2)$. As far as the number sequence is concerned, we can apply $h(n) = cn^{-\alpha}$, some positive α. Restrictions (2.4)-(2.5) are satisfied for $0 < \alpha < 1$, while (2.4) and (4.3) hold for $0 < \alpha < 1/2$.

Algorithms presented here are non-parametric, which means that the *a priori* knowledge of the identified system may be very small. The main feature of our approach, and advantage over parametric methods, is that our algorithms need no modification when the noise becomes correlated. Moreover, correlation of noise does not worsen convergence rate of the algorithms given in the paper (compare [2]-[5] where only white noise was admitted). It should be also emphasized that the proposed identification procedures, consisting generally in estimating regression, are simple from the computational viewpoint. An alternative approach, based on orthogonal series estimates of a regression function, has been presented in [6].

APPENDIX

The first lemma deals with the noise disturbing systems examined in the paper.

Lemma 1. We have

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\operatorname{cov}[Z_i, Z_j] = O(n).$$

The second lemma is of more general character.

Lemma 2. Let $E\{m(U_0)\} < \infty$. Let a Borel measurable kernel K satisfy (2.3). Then

$$(1/h)E\{m(U_0)K[(u-U_0)/h)]\} \to m(u)f(u)\int_{-\infty}^{\infty}K(v)dv \quad \text{as } h \to 0,$$

at every point u at which both f and m are continuous.

The lemma can be easily derived from Theorem 9.9 in [8, p.150].

REFERENCES

[1] J. S. Bendat, *"Nonlinear System Analysis and Identification"*, Wiley, New York, 1990.

[2] W. Greblicki, *"Nonparametric Estimation in Hammerstein Systems Based on Ordered Observations"*, IEEE Trans. Signal Proc., vol. **SP-44**, 1996 (in print).

[3] W. Greblicki, *"Nonparametric Identification of Wiener Systems by Orthogonal Series"*, IEEE Trans. Automat. Control, vol. **AC-39**, pp. 2077-2086, 1994.

[4] W. Greblicki, M. Pawlak, *"Nonparametric Identification of Hammerstein Systems"*, IEEE Trans. Inform. Theory, vol. **IT-35**, pp. 409-418, 1989.

[5] W. Greblicki, *"Nonparametric Identification of Wiener Systems"*, IEEE Trans. Inform. Theory, vol. **IT-38**, pp. 1487-1493, 1992.

[6] W. Greblicki, Z. Hasiewicz, *"Non-linear System Identification in the Presence of Correlated Noise"*, Technical Report, series **PRE-37/95**, Wrocław, 1995.

[7] G. Härdle, *"Applied Nonparametric Regression"*, Cambridge University Press, Cambridge, UK, 1990.

[8] R. L. Wheeden, A. Zygmund, *"Measure and Integral"*, Dekker, New York, 1977.